Development and Validation of Force-Field Parameters for Molecular Simulations of Peptides and Proteins Containing Open-Shell Residues

VINCENZO BARONE,[1] GABRIELLA CAPECCHI,[1] YVON BRUNEL,[2] MARIE-LOUISE DHEU ANDRIÉS[2], ROBERT SUBRA[2]

[1] *Dipartimento di Chimica, Università di Napoli ''Federico II,'' via Mezzocannone 4, I-80134 Napoli, Italy*

[2] *Laboratoire d'Etudes Dynamiques et Structurales de la Sélèctivité (LEDSS), Université Joseph Fourier, Grenoble, France*

Received 7 January 1997; accepted 18 April 1997

ABSTRACT: Parameters suitable for extending the AMBER force field for nucleic acids and proteins to open-shell derivatives of amino acid residues are proposed and tested. Two new atom types (radical carbon [CE] and hydrogen directly bonded to it [HE]) are introduced, whose parameters have been determined by a best fitting of quantum-mechanical computations of the simplest analogue of glycine radical (GlyR) in a peptide. The new force field is able to fit the reference results concerning both the structural parameters and the relative stabilities of the different conformers. It has been next applied to a conformational study of the distortions induced by extraction of the glycine H^{α} atom in an initially helical structure of a dodecamer of alanine including a central glycine residue. Our results show that the helical structure corresponds to a local energy minimum, but deeper minima are found which correspond to a fully planar GlyR residue included in a distorted helical sequence. © 1997 John Wiley & Sons, Inc. *J Comput Chem* **18**: 1720–1728, 1997

Keywords: force field; open-shell amino acids; peptides; glycine; glycine radical

Introduction

Numerical simulations have become, in recent years, one of the standard tools in the investigation of structural properties of biomolecules.[1–4] These techniques rest on the availability of simple and accurate force fields for the description of potential energy surfaces governing conformational transitions. Although experimental data can often be used to fit the parameters, the situation is more involved for unstable species like organic radicals. From an experimental point of view, biomolecules including free radical species are re-

Correspondence to: V. Barone

CCC 0192-8651 / 97 / 141720-09

ceiving increasing attention due to their role as degradation products during reaction with OH radicals or upon exposure to radiation.[5,6] It was then found that enzymes may specifically utilize radicals on intrinsic amino acid residues, the best known being the tyrosyl radical.[7] Glycine radicals,[8] formed during activation of two anaerobic enzymes, the pyruvate formate lyase (PFL)[9] and the ribonucleotide reductase (RNR),[10] have been identified recently by electron paramagnetic resonance (EPR) spectroscopy. Such radicals are not as well characterized as tyrosyl radicals and their conformational behavior is still under active experimental study.[9c] Recent post-Hartree–Fock computations[11] showed that the radical obtained upon abstraction of the H^{α} atom from a glycine dipeptide analogue has a conformational behavior quite different from that of dipeptide analogues of natural amino acidic species. As a consequence, the onset of regular structures for polypeptides containing the glycine radical cannot be studied until a specifically tailored force field has been built from reliable quantum-mechanical results.[12] As a first step in this direction we considered the model systems shown in Figure 1, namely a simple dipeptide analogue, and the radical derived from it upon abstraction of the H^{α} atom, both of which have already been studied by quantum-mechanical methods.[11,13] We have performed some further computations to complete the quantum-mechanical results, which provide the basis for the fitting of molecular mechanics parameters. Among the various force fields available, the last version of AMBER[14] appears to model accurately conformational energies and intermolecular interactions involving proteins, nucleic acids, and other molecules with related functional groups, which are of interest in biological chemistry. As a consequence, we have added to this force field two new atom types describing carbon radical centers and the hydrogen atoms directly bonded to them. Because dipeptides mimic at most local interactions in proteins, a better insight in the overall balance of intra- and interresidue noncovalent interactions can be obtained by the study of longer motifs. Consequently, we will analyze regular structures of oligomers of L-alanine, in which a single glycine or glycine radical residue has been inserted.

Methods

All the quantum-mechanical computations have been performed by the GAUSSIAN-94 package using the standard 6-31G* basis set.[15] Geometry optimizations have been performed at the unrestricted Hartree–Fock (UHF) level, with more reliable energetic values then being obtained by single-point UMP2 computations.

All molecular mechanics computations have been performed with the AMBER/95 force field available in version 2.95 of the Insight/Discover package.[16] Except for the new parameters described in the text, all the other parameters are the same as those described in Ref. 14.

Model compounds of the general form Ac—$(X)_n$—Y—$(X)_n$—NHMe have been considered in molecular mechanics computations, whereas only the simple F—Y—NH_2 model has been used in quantum-mechanical computations. Here Ac, NHMe, F, and NH_2 stand for acetyl, amidic $NHCH_3$, formyl, and amidic NH_2 groups, respectively; X is an alanyl residue, Y can be a glycyl (Gly) or glycyl radical (GlyR) residue, and n is 0 or 6. The appropriateness of formyl and NH_2 groups in place of the more common acetyl and $NHCH_3$ moieties has been tested in previous studies.[11–13]

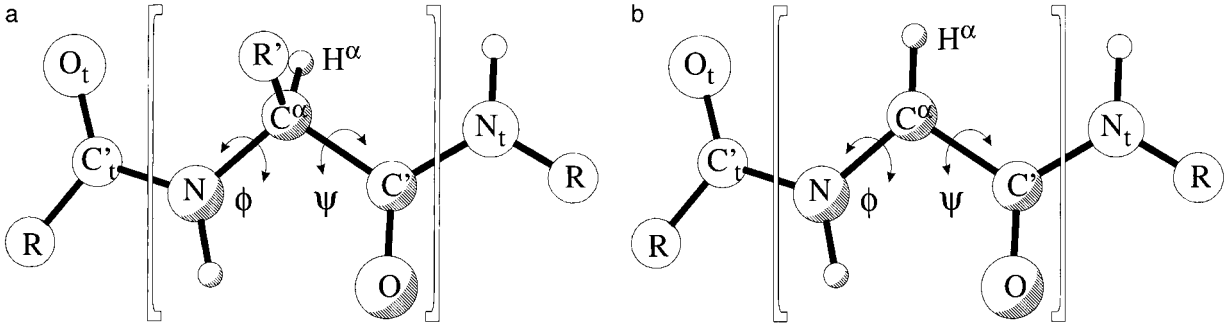

FIGURE 1. Atom labeling and principal geometric parameters of analogues of glycine (a) and glycine radical (b) in a peptide in their fully extended conformations.

All molecular mechanics minimizations were done with the complete Newton–Raphson minimizer and a convergency criterion of 0.001 kcal mol^{-1} $Å^{-1}$. Molecular dynamics simulations were carried out at a constant temperature of 500 K for a duration of 500 ps (the first 50 ps being excluded from the analysis), with a time step of 1.0 fs and a nonbonded cutoff of 9.5 Å. A representative number of structures resulting from the simulations were then selected and fully minimized.

The conformations of amino acid residues can be approximated by the values of the conventional torsional angles ϕ and ψ about the central carbon atom. Regular structures are further classified in terms of the number of atoms forming the ''cycle'' closed by an H-bond between NH and OC groups belonging to the backbone. According to this scheme, the extended conformation typical of β-sheets is labeled $C_5(\phi \cong 180°, \psi \cong 180°)$; γ turns are labeled $C_{7ax}(\phi \cong -90°, \psi \cong 60°)$ or $C_{7eq}(\phi \cong 90°, \psi \cong -60°)$. Conformers that could lead to helices in larger dipeptides (through the formation of C_{13} cycles) are labeled $\alpha_R(\phi \cong -60°, \psi \cong -30°)$ and $\alpha_L(\phi \cong 60°, \psi \cong 30°)$. Other conventional labelings are: P_{II} for a typical polyproline conformation ($\phi \cong -90°$, $\psi \cong 150°$), α' for semiextended structures ($\phi \cong 180°$, $\psi \cong 60°$), and β for bridge structures ($\phi \cong 90°$, $\psi \cong 0°$). Note that for Gly and GlyR residues both C_7 and α conformations are equivalent.

Results and Discussion

In the AMBER model the total energy is partitioned as follows:[14]

$$E_{total} = \sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\varphi - \gamma)] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{R_{ij}}\right]$$

Two new atom types (CE and HE) have been added to the standard atom list, whose van der Waals parameters have been set equal to those of the C and H1 atom types of AMBER. Fitting of the HF/6-31G* electrostatic potential for several conformers of F—GlyR—NH_2 shows that the charges of all the atoms of the GlyR moiety are quite different from the AMBER values for standard amino acid residues, and their variation as a function of ϕ and ψ angles is particularly strong due to the significant conjugation within the peptidic unit of this residue. We have, therefore, decided to use a minimalist approach, allowing a straightforward extension to radicals derived from H^α abstraction in other amino acid residues and/or to different force fields without further quantum-mechanical computations. This target is reached retaining the AMBER charges of the parent closed-shell residue for all the atoms except C^α, whose charge is univocally determined by the electroneutrality condition. In this model, the GlyR moiety can be obtained removing the whole side chain from any of the standard amino acids. Because the charge of H^α is near to 0.1 $|e^-|$ for most neutral amino acid residues,[14] we have used this value, which, in turn, leads to a charge of 0.0144 $|e^-|$ for C^α. We have verified that small modifications of these charges do not alter the structure and conformational behavior of the GlyR residue.

The other parameters involving CE and HE atoms are given in Table I and were obtained as follows. All the force constants for stretching and

TABLE I.
New Parameters Added to AMBER Force Field.

Bond	R_{eq} (Å)	K_r (kcal mol^{-1} $Å^{-2}$)
CE—HE	1.070	367.0
CE—N	1.377	367.0
CE—C	1.462	367.0
Angle	θ_{eq} (deg.)	K_θ (kcal mol^{-1} $deg.^{-2}$)
N—CE—C	114.0	63.0
N—CE—HE	119.0	35.0
C—CE—HE	127.0	34.0
CE—C—O	117.5	70.0
CE—C—N	118.0	80.0
CE—N—C	118.5	50.0
CE—N—H	118.0	30.0
Dihedral	$\frac{V_1}{2}$ (kcal mol^{-1})[a]	$\frac{V_2}{2}$ (kcal mol^{-1})[a]
X—N—CE—X	0.5	2.0
X—CE—C—X	0.5	2.0
N—CE—C—O	0.9	2.0
HE—CE—C—N	0.9	2.0

[a] $\gamma = 180°$.

bending are transferred without modifications from similar atom types already present in the AMBER force field. Equilibrium values for bond lengths and valence angles were initially set equal to the values obtained from the HF/6-31G* geometry optimization of the C_5 conformer of F—GlyR —NH_2. This set of parameters was already quite satisfactory for all the conformers, so that the final fine tuning of parameters could be done by ''manual adjustment'' without resorting to any specific algorithm. The energy differences between C_5, α', and β conformers completely determines the V_1 contributions for ϕ and ψ angles. Assuming next that the V_2 term is equal for both torsional angles, the best value of this constant is arrived at from a trial-and-error procedure on the relative energies of C_7 and helical structures. Note that only C_5, α' and β conformers correspond to true energy minima, whereas the geometries of the other conformers have been optimized by freezing the ϕ and ψ angles at their reference values.

As can be seen in Table II, there is good agreement between molecular mechanics and quantum-mechanical results, regarding the geometries and relative stabilities of true energy minima. This is also the case for the C_7[ΔE(QM) = 16.1 kcal mol^{-1}; ΔE(MM) = 16.6 kcal mol^{-1}] and α[ΔE(QM) = 15.8 kcal mol^{-1}; ΔE(MM) = 14.2 kcal mol^{-1}] conformers, which, although not corresponding to true energy minima, represent regions of the conformational space characteristic of natural peptides and proteins.

We next built a complete (ϕ, ψ) map of Ac—GlyR—NHMe using the rigid rotor (RR) approximation together with a standard set of bond lengths and valence angles[17] (the so-called Ramachandran map) and a resolution of 20°. Comparison with the corresponding quantum-mechanical map for F—GlyR—NH_2 (Fig. 2) provides a further validation of our parameters.

At this point, our results indicate that GlyR has a very rigid backbone structure, with energy minima corresponding to planar or nearly planar conformers. Such planarity fits naturally with one of the regular structures of peptides, namely the β-sheet one. However, most proteins are characterized by long helical regions stabilized by interresidue hydrogen bridges. It is, therefore, important to investigate the relative strengths of the intrinsic properisity to planarity of the GlyR residue on the one hand and the interresidue interactions on the other hand. To this end we performed a number of computations to compare the behavior of Gly and GlyR residues inserted in the middle of a polyalanine chain: Ac—$(Ala)_6$—Gly —$(Ala)_6$—NHMe (hereafter $Ala_{12}M$) and Ac—$(Ala)_6$—GlyR—$(Ala)_6$—NHMe (hereafter $Ala_{12}R$).

Starting from a regular α-helix, a full geometry optimization of $Ala_{12}M$ was performed. Then, one of the H^{α} atoms was eliminated and, after proper

TABLE II.
Geometrical Parameters (Bond Lengths in Angstroms and Angles in Degrees) and Relative Stabilities (ΔE and kcal mol^{-1}) Obtained by UHF/6-31G*//UMP2/6-31G* (QM) and Molecular Mechanics (MM) Computations.

	C5		α'		β	
Parameter	QM	MM	QM	MM	QM	MM
$C^{\alpha}H^{\alpha}$	1.070	1.071	1.069	1.071	1.073	1.070
NH	0.998	1.010	0.996	1.003	0.996	1.006
$C^{\alpha}N$	1.377	1.392	1.389	1.399	1.404	1.408
$C^{\alpha}C'$	1.462	1.473	1.453	1.475	1.475	1.485
$C'O$	1.215	1.231	1.193	1.226	1.205	1.228
$NC^{\alpha}H^{\alpha}$	118.2	118.6	118.3	118.2	115.2	113.6
$C'C^{\alpha}H^{\alpha}$	125.3	125.0	120.4	122.5	114.4	118.4
$NC^{\alpha}C'$	116.5	116.4	121.4	119.2	128.4	127.8
$C^{\alpha}NH$	116.4	117.2	118.4	121.1	115.4	114.5
$C^{\alpha}NC'_t$	123.6	122.5	122.8	121.6	128.4	132.6
$C^{\alpha}C'O$	119.6	118.9	112.7	117.1	117.0	116.5
$C^{\alpha}C'N_t$	116.3	119.4	116.7	120.8	119.9	120.2
ϕ	180.0	180.0	172.0	180.0	49.0	15.9
ψ	180.0	180.0	−10.0	0.0	−10.0	2.8
ΔE	0.0	0.0	6.9	6.7	7.9	7.7

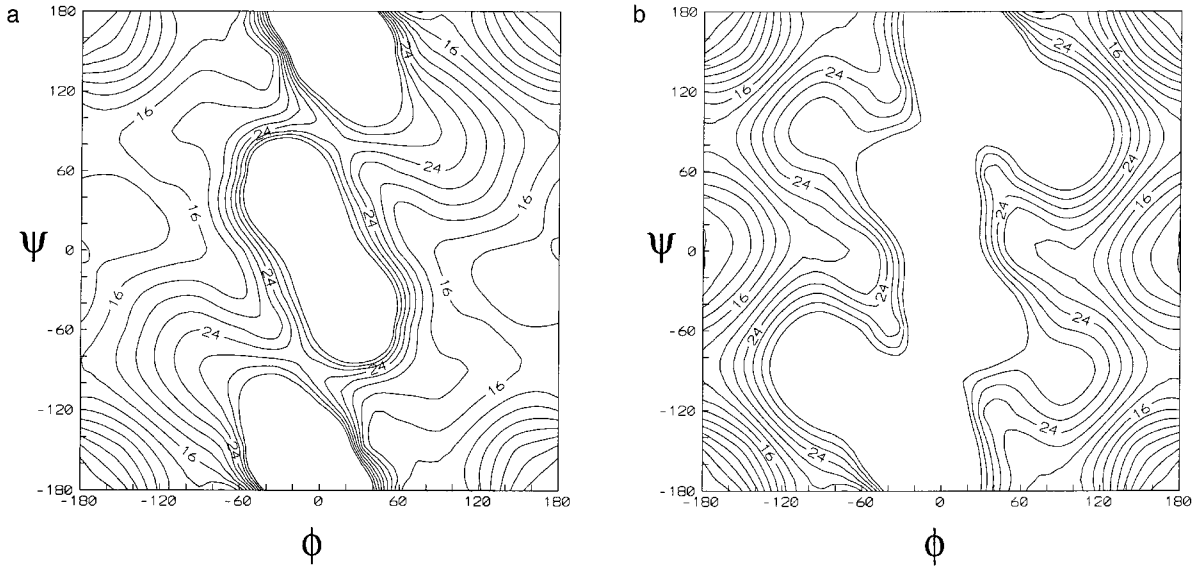

FIGURE 2. Ramachandran maps of F—GlyR—NH_2 computed at the HF / 6-31G* level (a) and of Ac—GlyR—NHMe computed with the new AMBER parameters (b). Contour lines are drawn every 2 kcal mol^{-1} up to 30 kcal mol^{-1} above the absolute minimum.

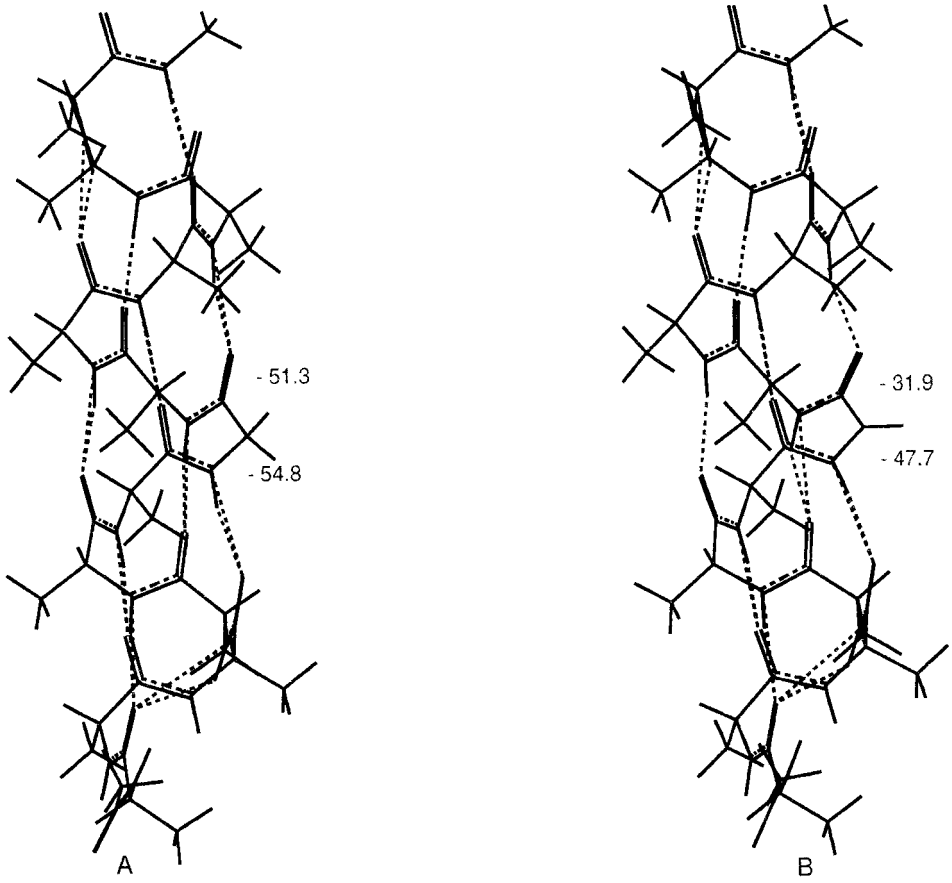

FIGURE 3. Perspective drawings of $Ala_{12}M$ (a) and $Ala_{12}R$ (b) in their optimized helical conformations evidencing the pattern of hydrogen bonds.

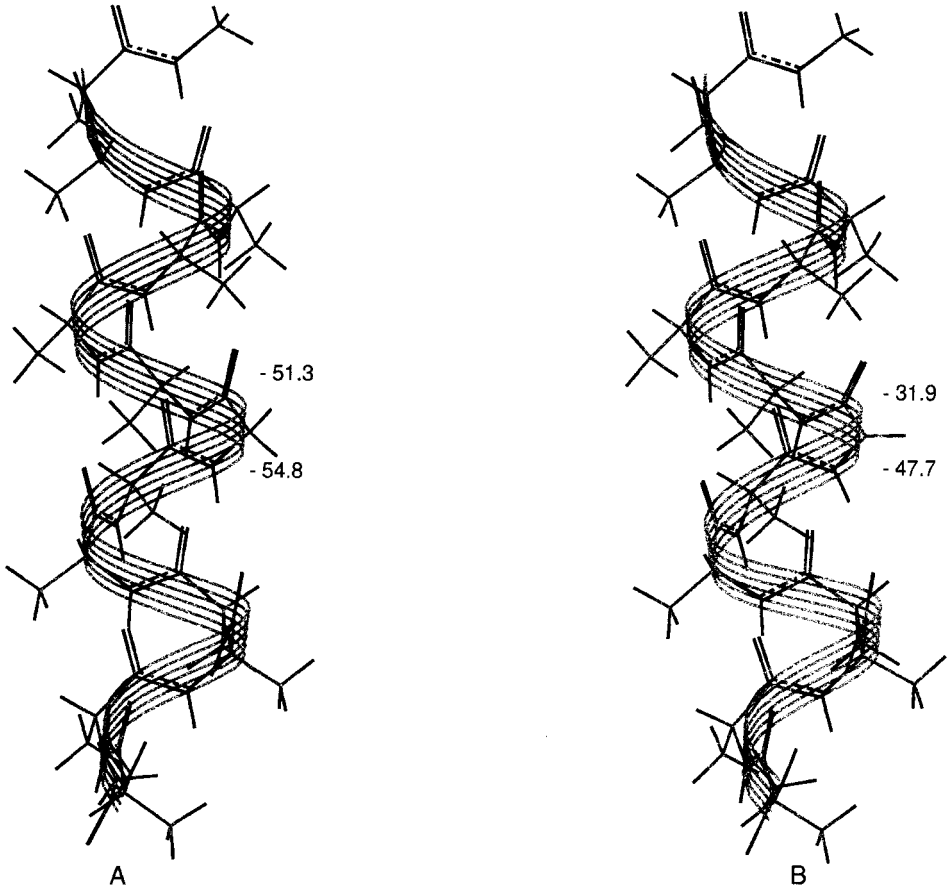

FIGURE 4. Perspective drawings of $Ala_{12}M$ (a) and $Ala_{12}R$ (b) showing the characteristics of their most stable helical structures.

modification of the atom types, a full geometry optimization of the resulting $Ala_{12}R$ structure was carried out. The minimized structures (Figs. 3 and 4) show that slightly distorted α-helices always correspond to local energy minima. Furthermore, it should be noticed that the formation of the radical modifies the ψ angle of the alanyl residue directly preceding it and the ϕ angle of the residue following it (Table III). This result is particularly significant for GlyR because it indicates that even strong intrinsic preferences of a residue cannot completely destroy the pattern of interresidue H-bonds dictated by the other residues. Starting from this structure, we performed a constant temperature molecular dynamics simulation at 500 K. It is apparent from Figure 5 that, after a relatively short period, the ϕ angle of the GlyR residue goes to a value near 180°, where it stays for the remaining

TABLE III.
Dihedral Angles of Central Unit and Its First Neighbors in Different Conformations of $Ala_{12}M$ and $Ala_{12}R$ (See Text).

	Ala_5		Ala_6		Y_7		Ala_8		Ala_9	
	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ
Y = Gly	−57°	−49°	−57°	−49°	−55°	−51°	−57°	−49°	−57°	−49°
Y = GlyR	−56°	−48°	−55°	−58°	−48°	−32°	−75°	−46°	−54°	−46°
Y = GlyR	−54°	−36°	−87°	165°	−178°	15°	53°	100°	−55°	−34°
Y = GlyR	−78°	−40°	83°	−39°	−179°	162°	−102°	28°	−77°	85°

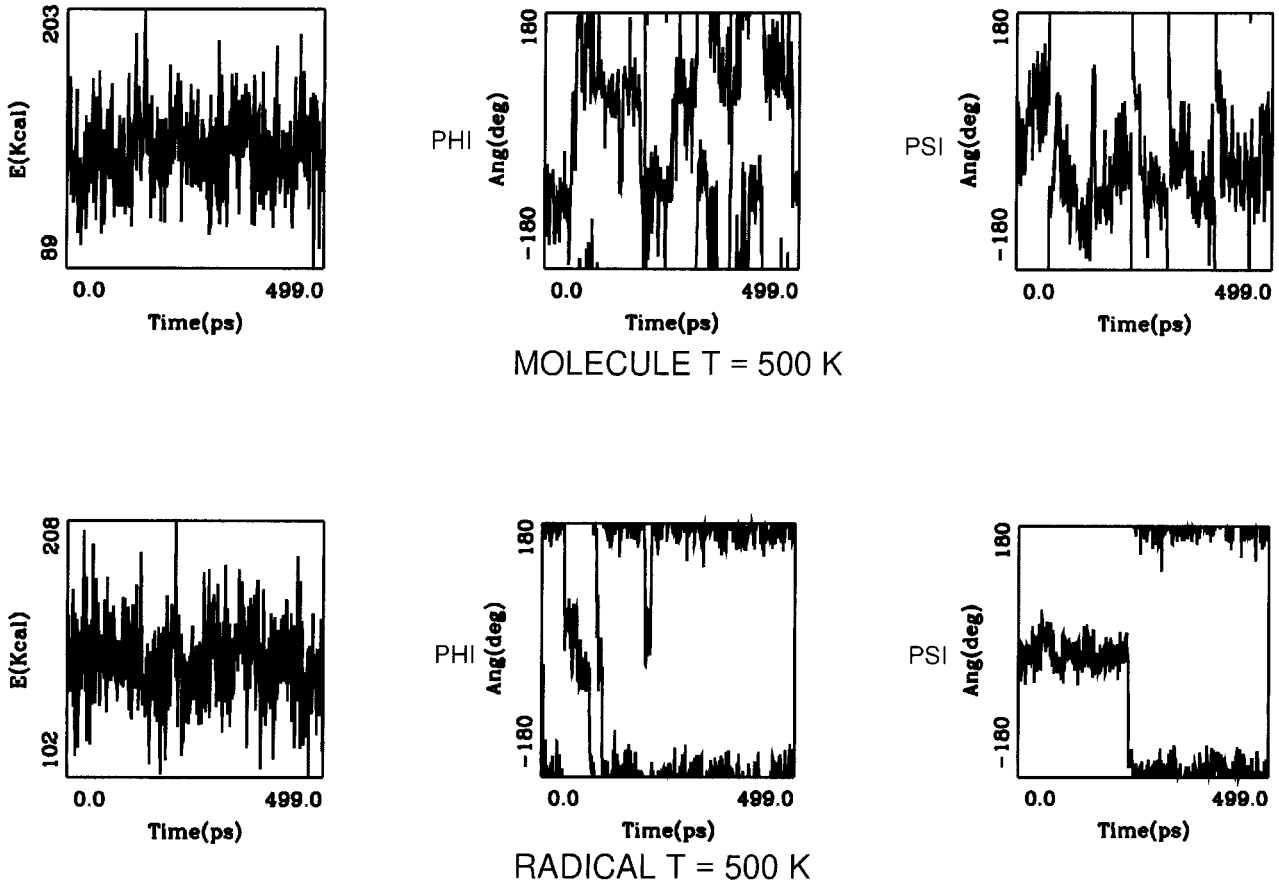

FIGURE 5. Evolution of the total energy and of the ϕ, ψ angles of the central residue obtained from constant temperature molecular dynamics simulations of $Ala_{12}M$ and $Ala_{12}R$ at 500 K.

time. The ψ angle shows the same general behavior, but it remains near 0° for a longer time before reaching its stable value of 180°. From this simulation it can be concluded that the GlyR residue retains its initial helical structure for a short period, followed by a conformational transition (involving a significant modification of the overall pattern of hydrogen bonds) leading to a GlyR residue in an α' conformation, and eventually (but after a longer time) the GlyR residue reaches it most stable C_5 conformation within a significantly distorted polypeptide. We next randomly selected four structures involving the GlyR residue in the α' conformation and four structures involving the GlyR residue in the C_5 conformation and performed full geometry optimizations. The final structures are shown, together with their relative energies, in Figure 6. It is quite apparent that one of the structures involving the GlyR residue in a C_5 conformation is the most stable, so we can suggest that, if sufficient time remains, a planar conformation of this residue becomes compatible with a favorable arrangement of interresidue hydrogen bridges, allowing the development of helical structures, albeit with significant distortions of their regular shape.

Conclusion

This article is devoted to the development and validation of an extension of the AMBER force field to open-shell amino acid derivatives. From a methodological point of view, our first results show that a limited number of additional parameters is sufficient to closely match the reference quantum-mechanical results for a simple dipeptide analogue. From a more chemical point of view, we have confirmed that the conformational freedom of open-shell amino acid residues is severely reduced with respect to that of their parent closed-shell systems due to the strong preference for planar or nearly planar conformations induced by a more effective electron conjugation along the

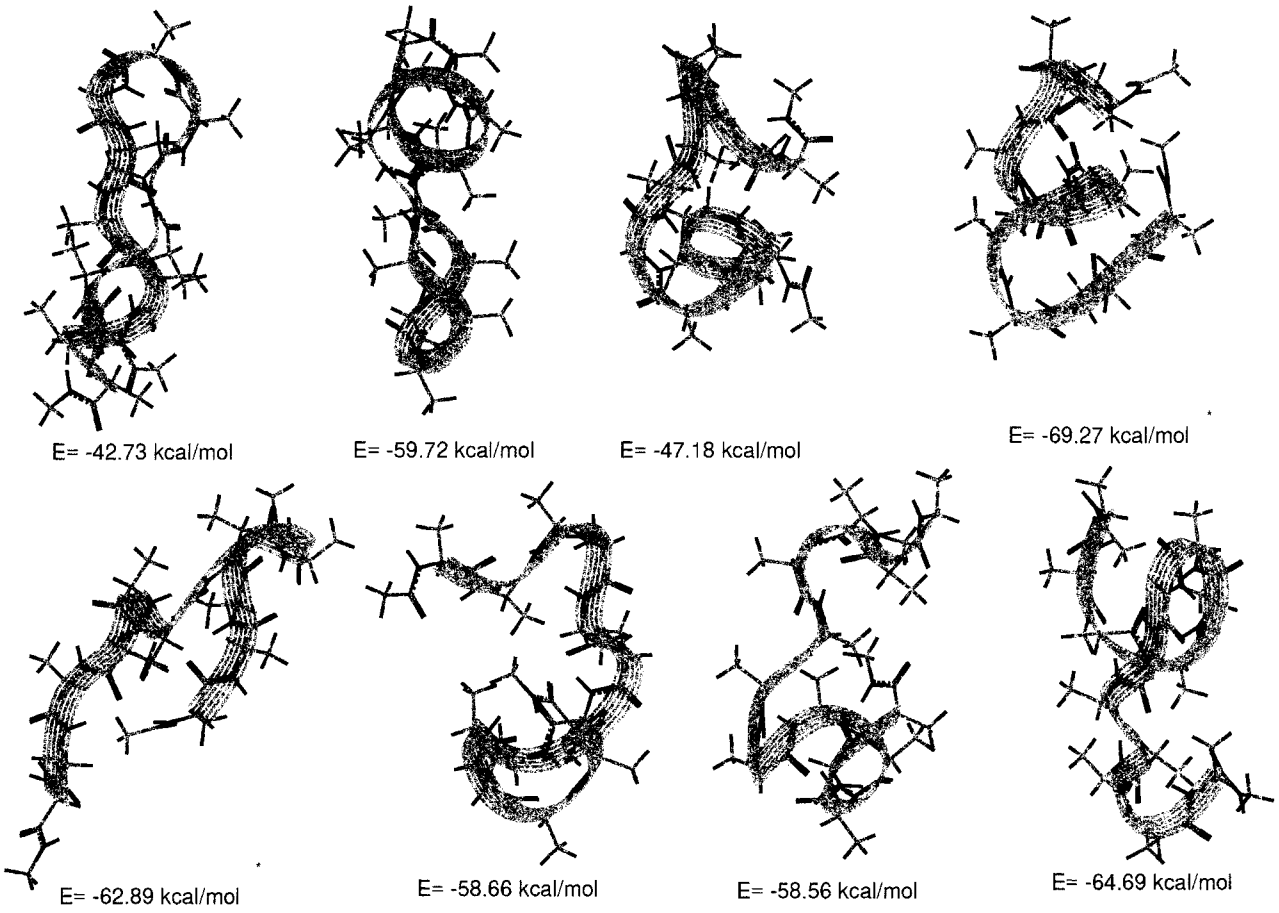

FIGURE 6. Representative optimized structures of $Ala_{12}R$ when the central GlyR residue is in a C_5 (upper part) of α' (lower part) conformation.

backbone. This has a significant effect on the behavior of relatively long peptides, whose structure is determined by a balance between local effects and the tendency to form interresidue hydrogen bonds. Work is in progress in our laboratories to extend these results to other residues and to investigate the relation between structure and spectroscopic properties of naturally occurring systems, including open-shell species.

References

1. S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, *J. Comput. Chem.*, **7**, 230 (1986).
2. W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988).
3. J.-H. Lii and N. L. Allinger, *J. Comput. Chem.*, **12**, 186 (1991).
4. J. R. Maple, M.-J. Hwang, T. P. Stockfisch, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler, *J. Comput. Chem.*, **15**, 162 (1994).
5. P. Neta and R. W. Fessenden, *J. Phys. Chem.*, **75**, 738 (1971).
6. R. Livingston, D. G. Doherty, and H. Zelders, *J. Am. Chem. Soc.* **97**, 3198 (1975).
7. (a) P. Reichard and A. Ehrenberg, *Science*, **221**, 514 (1983); (b) M. Sahlin, L. Petersson, A. Gräslund, A. Ehrenberg, B. M. Sjöberg, and L. Thelander, *Biochemstry*, **26**, 5541 (1987); (c) C. Bender, M. Sahlin, G. T. Babcock, A. B. Barry, T. K. Chandrasekhar, S. P. Salowe, J. Stubbe, B. Lindstrom, L. Petersson, A. Ehrenberg, and B. M. Sjoberg, *J. Am. Chem. Soc.*, **111**, 8079 (1989); (d) J. A. DeGray, G. Lassmann, J. F. Curtis, T. A. Kennedy, L. A. Marnett, T. E. Eling, and R. P. Mason, *J. Biol. Chem.*, **267**, 23583 (1992); (e) G. T. Babcock, M. K. El-Deeb, P. O. Sanduski, M. M. Whittaker, and J. W. Whittaker, *J. Am. Chem. Soc.*, **114**, 3727 (1992); (f) C. W. Hoganson and G. T. Babcock, *Biochemistry*, **31**, 11874 (1992).
8. (a) E. Hayon and M. Simic, *J. Am. Chem. Soc.*, **93**, 6781 (1971); (b) Y. Kirino and H. Taniguchi, *J. Am. Chem. Soc.*, **98**, 5098 (1976).
9. (a) V. Unkrig, F. A. Neugebauer, and J. Knappe, *Eur. J. Biochem.*, **184**, 723 (1989); (b) A. F. V. Wagner, M. Frey, F. A. Neugebauer, W. Schäfer, and J. Knappe, *Proc. Natl. Acad. Sci. USA*, **89**, 996 (1992); (c) M. Frey, M. Rothe, A. F. V. Wagner, and J. Knappe, *J. Biol. Chem.*, **269**, 12432 (1994).
10. E. Mulliez, M. Fontecave, J. Gaillard, and P. J. Reichard, *J. Biol. Chem.*, **268**, 2296 (1993).

11. V. Barone, C. Adamo, A. Grand, Y. Brunel, M. Fontecave, and R. Subra, *J. Am. Chem. Soc.*, **117**, 1083 (1995).
12. P. Amodeo and V. Barone, *J. Am. Chem. Soc.*, **114**, 9085 (1992).
13. (a) T. Head-Gordon, M. Head-Gordon, M. J. Frisch, C. L. Brooks III, and J. A. Pople, *J. Am. Chem. Soc.*, **113**, 5989 (1991); (b) L. H. S. Shang and T. Head-Gordon, *J. Am. Chem. Soc.*, **116** 1528 (1994).
14. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Sellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179 (1995).
15. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. DeFrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, GAUSSIAN-94, Rev. D.4, Gaussian, Inc., Pittsburgh, PA, 1995.
16. Insight/Discover is available from Biosym/MSI, 9685 Scranton Road, San Diego, CA 92121-2777.
17. E. Benedetti, In *Peptides*, M. Goodman and J. Meinhofer, Eds., John Wiley & Sons, New York, 1977, p. 257.